

Nuthall, Graham A.

Ausgewählte neuere Untersuchungen zur Unterrichtsinteraktion und zum Lehrverhalten

Wulf, Christoph [Hrsg.]: Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München : Piper 1972, S. 239-263



Quellenangabe/ Reference:

Nuthall, Graham A.: Ausgewählte neuere Untersuchungen zur Unterrichtsinteraktion und zum Lehrverhalten - In: Wulf, Christoph [Hrsg.]: Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München : Piper 1972, S. 239-263 - URN: urn:nbn:de:01111-opus-14294 - DOI: 10.25656/01:1429

<https://nbn-resolving.org/urn:nbn:de:01111-opus-14294>

<https://doi.org/10.25656/01:1429>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, veröffentlichen oder widernatürlich nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Evaluation

Beschreibung und Bewertung von Unterricht,
Curricula und Schulversuchen

Texte

herausgegeben von Christoph Wulf



R. Piper & Co. Verlag
München

ISBN 3-492-01985-4
© R. Piper & Co. Verlag, München 1972
Gesamtherstellung Clausen & Bosse, Leck/Schleswig
Umschlagentwurf Gerhard M. Hotop
Printed in Germany

GRAHAM A. NUTHALL¹

*Ausgewählte neuere Untersuchungen
zur Unterrichtsinteraktion und zum Lehrverhalten*

Ein kritischer Bericht

Den meisten Lesern wird aufgefallen sein, daß in letzter Zeit die Zahl der Untersuchungen über das Lehrerverhalten und die Lehrer-Schüler-Interaktion im Unterricht zugenommen hat. Einige neuere Untersuchungen machen das Ausmaß dieser Zunahme deutlich (Lawrence 1966; Amidon/Simon 1965; Gage/Unruh 1967; Biddle 1967). Obwohl derartige Untersuchungen nicht neu sind, haben viele Studien noch immer Versuchscharakter. Daher wurden bisher nur wenige von ihnen in größerer Auflage veröffentlicht; und es wurde selten versucht, die bisherigen oder noch erreichbaren Ergebnisse dieser neuen Forschungsrichtung kritisch einzuschätzen. Dieser Beitrag soll einen kritischen Überblick über eine Auswahl neuerer Untersuchungen geben, die für diesen Forschungszweig sehr wichtig zu sein scheinen. Besondere Aufmerksamkeit soll dabei jenen Befunden zukommen, die unsere Kenntnis über die Beziehungen zwischen den Verhaltensweisen des Lehrers und den Lernprozessen bzw. Leistungen des Schülers verbessern.

Von ein paar Ausnahmen abgesehen, wurde das Lehrerverhalten und die Unterrichtsinteraktion nach dem klassischen Modell aller wissenschaftlichen Untersuchungen erforscht. Nach einer Reihe erkundender und abtastender Felduntersuchungen führten gezielte Versuche, Taxonomien des Unterrichtsverhaltens zu erarbeiten, zu mehreren hochentwickelten deskriptiven Untersuchungen über Lehrer und Schüler *in situ*. Im Anschluß an diese ersten »ökologischen« Überblicke wurde dann versucht, die beobachteten Lehrvariablen mit der gemessenen Schülerleistung zu korrelieren. Und neuerdings gibt es auch sorgfältig kontrollierte Untersuchungen halbexperimenteller Art. Da die Wirklichkeit jedoch stets verworrener ist als das Modell, finden immer noch interessante, wichtige Entwicklungen auf jeder der verschiedenen Entwicklungsstufen statt.

Daher werden die Untersuchungen, die analysiert werden sollen, unter drei Rubriken betrachtet: Zuerst soll die Entwicklung der Methoden behandelt werden, die zur Beobachtung und Analyse des gesamten Unter-

richtsgeschehens dienen. Dabei wird nicht versucht, alle der mehr als dreißig verschiedenen Systeme zu behandeln; es sollen lediglich einige interessante neue Verfahren sowie einige Modifikationen alter Systeme beschrieben werden. Danach sollen einige Ergebnisse der Unterrichtsbeobachtung dargestellt werden; dabei soll besonders berücksichtigt werden, wie weit sie zu unserem Verständnis der Ursachen des Lernens beitragen. Im letzten Teil werden dann Untersuchungen behandelt, in denen versucht wird, Schülerleistung in Beziehung zu Lehrerverhalten oder Unterrichtsinteraktion zu bringen. Sie reichen von einfachen Korrelationsuntersuchungen der Beziehungen in normalen Klassen, die ihren üblichen Unterricht erhielten, bis hin zu Versuchen, die Lehrvariablen experimentell zu manipulieren.

Verfahren zur Analyse von Unterrichtsverhalten

Revision eines alten Systems

Das bei weitem am häufigsten gebrauchte Verfahren für die Analyse des Unterrichtsverhaltens ist das Flanderssche Verfahren der Interaktionsanalyse (Amidon/Flanders 1963). Nach diesem Verfahren erstellt ein Beobachter ein fortlaufendes Protokoll der Unterrichts-Aktivitäten von Lehrer und Schülern. Alle drei Sekunden (bei wichtigen Veränderungen noch öfter) notiert der Beobachter die Kategorie, die das augenblickliche Geschehen am besten beschreibt. Das Ergebnis ist ein kontinuierliches Protokoll, das in bezug auf die relative Häufigkeit verschiedener Verhaltensarten und hinsichtlich der paarweisen Verhaltensfolgen anhand einer zwei-dimensionalen Matrix analysiert werden kann. Der ursprüngliche Zweck der Interaktionsanalyse bestand darin, Schätzwerte zu erhalten, bis zu welchem Grad die Lehrer ihre Schüler »indirekt« oder »direkt« beeinflussen. Das ursprüngliche, von dieser Analyse abgeleitete Maß (I/D Verhältnis) sollte diesen Aspekt des »affektiven Klimas« (affective climate) im Unterricht ausdrücken.

Während dieses Verfahren in seiner ursprünglichen Form noch immer sehr häufig benutzt wird, haben auch mehrere Verbesserungsversuche stattgefunden. In seinem Bericht über neue Entwicklungen bei der Interaktionsanalyse wies Amidon (1966) auf das Fehlen »kognitiver« Kategorien hin. Er beschrieb eine Reihe von Kategorien zur Identifikation verschiedener kognitiver Ebenen bei Fragen und Antworten. Teilweise basierten sie auf der Arbeit Aschners und Gallaghers (Aschner/Gallagher u. a. 1965), die auf Guilfords fünf kognitive Prozesse zurückgreift (Gedächtnis, Er-

kenntnis [cognition], konvergierende und divergierende Produktion und Bewertung). Bisher gibt es offenbar noch keine Untersuchungen, die eine Modifikation des Flandersschen Verfahrens in bezug auf kognitive Variablen benutzen; einige empirische Untersuchungen deuten darauf hin (Furst 1967), daß das Konzept des »kognitiven Klimas« im Unterricht weiter ausgearbeitet werden sollte. Ein solcher Meßwert des »kognitiven Klimas« im Unterricht könnte zeigen, welches »kognitive Niveau« (cognitive level) für einen Lehrer charakteristisch ist. L. Siegel und L. C. Siegel (1967) entwickelten ein ähnliches Konzept für das »intellektuelle Klima« (intellectual climate).

Eine weitere Modifikation des Flandersschen Verfahrens, typisches Lehrerverhalten zu messen, versuchte Honigman (1968). Er berichtete über die Entwicklung eines dreidimensionalen Kategorien-Systems und fügte der ursprünglichen affektiven Dimension eine kognitiv-inhaltliche (cognitive-substantive) sowie eine Verfahrensdimension (class management) hinzu. Dieses System wird als »Multidimensionale Analyse der Unterrichtsinteraktion« bezeichnet (Multidimensional Analysis of Classroom Interaction; MACI). Dieser Versuch, die Skala der Verhaltensweisen, auf die der Beobachter zu achten hat, auszuweiten, hat Honigman offenbar dazu gebracht, einige einfallsreiche Aufzeichnungstechniken zu entwickeln, die das Protokollieren erleichtern sollen. Honigmans Ausführungen sind auch bemerkenswert, weil sie einen Versuch beschreiben, die Validität (concurrent validity) der »Multidimensionalen Analyse der Unterrichtsinteraktion« zu prüfen. Bei dieser Untersuchung wurden 75 Lehrer, die an einem Sommerkurs teilnahmen, nach einem Zufallsstichprobenverfahren in drei gleich große Gruppen geteilt. Die eine Gruppe sah die Fernsehaufzeichnung einer Unterrichtsstunde, die zweite las nur das nach der »Multidimensionalen Analyse der Unterrichtsinteraktion« von einem Beobachter geführte Protokoll derselben Stunde, und die dritte Gruppe las ein Protokoll, das nach dem Flandersschen Verfahren geführt worden war. Der Kriteriumswert (criterion measure) wurde auf Grund der Angaben aller drei Gruppen in einem Fragebogen ermittelt, der aus 32 Fragen über verschiedene (affektive, kognitive und Verfahrens-)Aspekte der Stunde bestand. Folgende Hypothese wurde aufgestellt: Die Antworten der Gruppe, die das nach Honigmans »Multidimensionaler Analyse der Unterrichtsinteraktion« angefertigte Protokoll gelesen hatte, mußten mehr Ähnlichkeiten mit den Antworten der Gruppe haben, die die Life-Übertragung der Stunde gesehen hatte, als mit denen der Gruppe, die das nach dem Flandersschen Kategorien-System angefertigte Protokoll gelesen hatte. Ein solches Ergebnis war bei dem Teil des Fragebogens bestimmt zu erwarten, der sich auf die kognitive Dimension der Unterrichtsstunde bezog. Die Er-

gebnisse zeigten jedoch, daß die »Multidimensionale Analyse der Unterrichtsinteraktion« zwar besser die affektiven Aspekte, nicht aber die kognitiven und die Verfahrens-Aspekte der Unterrichtsstunde herausarbeitete. Es konnte noch nicht nachgewiesen werden, daß Honigmans Beobachtungssystem wirklich multidimensional ist.

Die Entwicklung neuer Systeme

Während nach dem Flandersschen Verfahren der Interaktionsanalyse und Honigmans »Multidimensionaler Analyse der Unterrichtsinteraktion« der Beobachter seine Aufzeichnungen in gleichmäßigen Zeitabständen machen muß, gibt es andere Analyse-Systeme, nach denen Beobachtungen in natürlichen Zeiteinheiten oder zu natürlichen Übergängen im Unterrichtsverhalten gemacht werden. Weil die nötige Identifikation und Analyse dieser Zeiteinheiten oder Übergangspunkte oft recht schwierig ist, werden in der Regel Tonband- oder Videorecorder-Aufzeichnungen des untersuchten Verhaltens gemacht. Bellack und andere (1966) berichteten z. B. von einer genauen Analyse der von ihnen als »Unterrichtszyklus (teaching cycles) bezeichneten Einheit. Die Analyse beruht auf Transkriptionen von Tonbandaufnahmen mehrerer Unterrichtsstunden, die über das gleiche ökonomische Thema von verschiedenen Lehrern gehalten wurden. Ein »Unterrichtszyklus« ist von allen natürlichen Einheiten im Ablauf der Diskussion im Unterricht am leichtesten als Einheit zu identifizieren; er besteht in der »Frage-Antwort-Reaktions-Sequenz«, oder in einer ihrer vielen Varianten. Dies war auch die Grundeinheit einer Untersuchung von Nuthall und Lawrence (1965).

In zwei sehr verschiedenen Untersuchungen wird von der Entwicklung neuer analytischer Systeme berichtet, die auf natürlichen Einheiten beruhen. Die erste Untersuchung beschreibt den Versuch, mit Hilfe eines brauchbaren psychologischen Konzepts das Unterrichtsverhalten zu erforschen. In der anderen Untersuchung wird der weit komplexere Versuch gemacht, Methoden zur Aufzeichnung aller denkbaren wichtigen Verhaltensweisen im Unterricht zu finden. MacDonald und Zaret (1967) berichteten von einem Versuch, den Begriff der »Offenheit« in den menschlichen Beziehungen (wie er von Rogers und anderen in der Psychotherapie entwickelt wurde) auf die Interaktion im Unterricht zu beziehen. Sie entwickelten ein Verfahren, den Grad an Offenheit in den protokollierten verbalen Reaktionen der Lehrer gerade zu den Zeitpunkten im Unterrichtsgespräch einzuschätzen, zu denen die Lehrer Antworten von Schülern kommentieren und die weitere Diskussionsrichtung bestimmen. Sie stellten die Hypothese auf, daß die Lehrer, die sich während dieser wichtigen Situati-

on in der Unterrichtsdiskussion offener verhielten, mehr produktive Antworten ihrer Schüler provozieren würden. Die Lehrer dagegen, die verschlossener sind oder sich stärker an den Rollenerwartungen orientieren, erhalten dadurch wahrscheinlich weniger produktive Antworten von ihren Schülern und erhöhen gleichzeitig auch die Zahl der nur reproduktiven Antworten. Durch Tonbandaufzeichnungen von Unterrichtsstunden in der Sozialkunde (social studies) in neun Klassen einer Elementarschule fanden sie in acht der neun Klassen statistisch signifikante Beziehungen zwischen dem Grad der Offenheit des Lehrers und der Zahl der produktiven Antworten der Schüler. Obwohl sicherlich noch mehr sorgfältig kontrollierte Daten nötig sind, scheint doch der Wert des Begriffs der Offenheit für die Unterrichtsinteraktion nachgewiesen worden zu sein. Es wäre höchst interessant zu wissen, ob es zwischen Flanders' Begriff der »Indirektheit« und MacDonalds und Zarets Begriff der »Offenheit« eine enge Beziehung gibt.

In ihrem Bericht über ihre umfangreiche Untersuchung über Unterrichtsaktivitäten beschrieben Biddle und Adams (1967) die Entwicklung eines Beobachtungssystems, das wohl das umfangreichste aller vorliegenden Systeme ist. Im Gegensatz zu dem System von MacDonald und Zaret beruht das System von Biddle und Adams nicht auf einer spezifischen Theorie oder einem theoretischen Konstrukt, obwohl es etwas soziologisch orientiert ist. Es scheint der wohlüberlegte Versuch zu sein, die Voraussetzungen dafür zu schaffen, daß möglichst viele verschiedene, leicht identifizierbare Aspekte aller Aktivitäten im Unterricht in einem Protokoll aufgenommen werden können. Der Umfang des Systems hängt offenbar davon ab, ob eine Videorecorder-Anlage vorhanden ist, die eine detaillierte Analyse verbaler und verhaltensbezogener Aspekte des Unterrichtsgeschehens ermöglicht. Das System von Biddle und Adams basiert auf der Unterscheidung zwischen strukturellen und funktionellen Aspekten der Unterrichtsaktivitäten. Der strukturelle Aspekt umfaßt Positionen und Rollen aller Unterrichtsteilnehmer (was sie tun, mit wem, wo, usw.). Der funktionelle Aspekt umfaßt die Art des Inhalts (oder der Bedeutung) der Interaktion und die Art der Behandlung dieses Inhalts. Um also Strukturen beschreiben zu können, wird identifiziert: Wer spricht (Sender), zu wem (Empfänger), und wer hört zu (Zuhörer), also die Kommunikationsstruktur und Gruppierung aller Individuen und das Verhältnis der Individuen zu den physikalischen Dimensionen der Klasse. Um Funktionen zu beschreiben, identifizieren sie die drei in Honigmans »Multidimensionaler Analyse der Unterrichtsinteraktion« enthaltenen Dimensionen, und differenzieren weiter zwischen relevanter und irrelevanter Thematik. Eine Sonderkategorie erfaßt nicht-sprachliche Vorgänge (z. B. Drill, praktische Übungen, usw.).

Im allgemeinen hat man bisher dazu geneigt, möglichst viele verschiedene Aspekte des Unterrichtsverhaltens zu erfassen, ohne die feineren Unterscheidungen und Unterordnungen zu berücksichtigen, die unter jedem dieser Aspekte vielleicht zu treffen wären. Das Ergebnis ist ein System, das in gewissem Sinne unbegrenzt erweiterbar ist, vorausgesetzt, daß die Video-recorder-Aufnahmen der Unterrichtsvorgänge ausreichend präzise sind.

Die Zukunft der Systeme zur Analyse der Unterrichtsinteraktion

Während der frühen Stadien der Entwicklung der ersten wichtigen Systeme zur Analyse des Unterrichtsverhaltens äußerten mehrere Autoren die Befürchtung, die Vielzahl verschiedener Systeme könne zu einer beträchtlichen Verwirrung führen. Wenn jeder sein eigenes Analyse-System aufbaue, so wurde befürchtet, werde eine gewaltige Informationssammlung über Unterrichtsverhalten zustande kommen, die nicht systematisiert werden könne und aus der noch nicht einmal die einfachsten Thesen über Unterrichtsgeschehen entwickelt werden können, die die meisten Forscher benötigen. Schon jetzt könne man die Situation verwirrend finden und vielleicht den Wert so vieler scheinbarer Wiederholungen in Frage stellen. Dazu führt Komisar aus:

Die Produktion »neuer« Kategoriensysteme bringt uns rasch dem Chaos näher. Scheinbar ist keiner der Forscher willens oder in der Lage, uns zu erklären, *warum* gewisse Kategorien gewählt werden oder in welchem Verhältnis die Kategorien eines Forschers zu denen eines anderen stehen (Komisar 1968, 22).

Andererseits wies Biddle darauf hin, daß jeder Forscher unbedingt sein eigenes System entwickeln müsse, um den Bedingungen seines eigenen Forschungsprojekts gerecht zu werden, zumal so vielen der neuen Systeme jegliche zuverlässige oder systematisch-theoretische Grundlage fehle (Biddle 1968, 31).

Mit dem Erscheinen einer Anthologie der 26 »bekanntesten und gebräuchlichsten Beobachtungssysteme für den Unterricht« (Simon/Boyer 1967) ist hoffentlich ein Wendepunkt erreicht worden². In Zukunft wird der Konstrukteur eines »neuen« Kategoriensystems sein Werk theoretisch rechtfertigen müssen. So wird vielleicht die nötige Aufmerksamkeit auf die Erklärung und das Verständnis des Unterrichtsverhaltens gelenkt. Wie in den folgenden Abschnitten dieses Beitrags gezeigt werden soll, ist trotz der vielen protokollierten Details über die Geschehnisse in zahlreichen Unterrichtssituationen immer noch sehr wenig über die Ursachen oder Wirkungen der Geschehnisse bekannt.

Beziehungen zwischen Beobachtungssystemen

Es wurde bereits mehrmals versucht, die Beziehungen zwischen verschiedenen Systemen zu erforschen. In einigen Untersuchungen wandte man dazu beispielsweise zwei verschiedene Systeme parallel an. Medley und Hill (1968) berichteten über die Ergebnisse eines Vergleichs des Flandersschen Verfahrens der Interaktionsanalyse mit ihrer eigenen, unlängst abgeänderten Fassung des »Beobachtungsplans und -protokolls« (Observation Schedule and Record; OSCAR 4V). Bei dieser Untersuchung wurden 70 Lehrer und Lehrerinnen der Sekundarschule während ihres ersten Berufsjahres von je zwei Beobachtern mindestens viermal systematisch beobachtet. Jedesmal protokollierte ein Beobachter nach dem Flandersschen System und der andere nach OSCAR 4V. Entsprechend diesen parallelen Protokollen wurden für jeden Lehrer insgesamt 75 Meßwerte interkorreliert (38 nach Flanders und 37 nach OSCAR). Da es offenbar schwierig ist, in eine so große Korrelationstabelle (75×75) Sinn hineinzubringen, wurde eine Faktorenanalyse durchgeführt. Es gelang, zehn Faktoren zu identifizieren, die die Unterschiede zwischen den Lehrern beschrieben. Von diesen zehn Faktoren wurden fünf mit beiden Systemen gemessen, drei nur mit OSCAR, und zwei nur mit dem Flandersschen System. Die beiden Systeme scheinen sich bis auf einige unabhängige Merkmale zu überschneiden. Es ist allerdings sehr schwierig, die Besonderheit jedes Systems aus den untersuchten Faktoren abzuleiten.

In einer von Pearson³ geleiteten Untersuchung wurde das Flanderssche Verfahren als zusätzliches Verfahren zur Analyse der »evaluativen Abschnitte« (evaluative ventures) bei Gesprächen in Klassen der Sekundarstufe I (intermediate school) benutzt. Ein »evaluativer Abschnitt« ist einer von neun Abschnitten oder themen-zentrierten Gesprächseinheiten (topic-centered units), die Smith und Meux (1967) identifiziert haben, unter dem die Diskussion über eine Wertfrage oder die Bedeutung eines Sachverhalts verstanden wird. Die vorläufige Analyse der Daten weist darauf hin, daß die Meßwerte des Flandersschen Systems mit der Zahl der evaluativen Gespräche, mit ihrer Dauer und ihrer relativ logischen Komplexheit in Beziehung stehen. Mit anderen Worten: Das »affektive Klima« ist erwartungsgemäß nicht unabhängig von den logischen und semantischen Aspekten der Diskussion in einer Schulklasse.

Durch Beobachtungssysteme gewonnene Ergebnisse

Einige neuere Untersuchungen brachten Daten über die Art des Lehrverhaltens und der Unterrichtsinteraktion; zu ihnen gehören auch die Untersuchungen von Bellack u. a. (1966) sowie von Smith und Meux (1967), deren Ergebnisse sich auf umfassende Forschungen stützen, die hauptsächlich verbale Interaktionen, semantische und thematisch-inhaltliche Aspekte betreffen.

Die Sequenz der Unterrichtsinhalte

Smith und Meux (1967) berichteten über die Ergebnisse eines Versuchs, die Organisation und sequenzielle Anordnung der Unterrichtsinhalte mit Hilfe von Aufzeichnungen von Diskussionen in Sekundarschulen zu analysieren. Sie definierten zunächst den Begriff »Unterrichtsstrategie« (teaching strategy) als eine besondere Sequenz von Teilthemen innerhalb des vorgegebenen Unterrichtsthemas. Davon ausgehend analysierten und beschrieben sie die »Strategien« in acht verschiedenen, zuvor identifizierten Kategorien. Diese acht Kategorien waren folgende: Ursachen, Begriffe, Bewertungen, Interpretationen, Verfahren, Regeln, Gründe und besondere Informationen (causes, concepts, evaluations, interpretations, procedures, rules, reasons and particular information). Dadurch fanden Smith und Meux z. B. heraus, daß Diskussionen über Begriffe zu der häufigsten Art von Diskussionen gehören. Nachdem sie diejenigen Teile des Gesprächs isoliert hatten, in denen ein genannter Begriff Mittelpunkt der Diskussion war, isolierten sie auch die Informationen über Begriffe, in denen die Begriffe ganz allgemein beschrieben oder besprochen wurden. Die kurzen Abschnitte der Diskussion, in denen nur eine einzige Information über einen Begriff beschrieben oder besprochen wurde, bezeichneten sie als »Begriffs-Impulse« (conceptual moves). So betrachtet, besteht der Begriff Unterrichtsstrategie also aus einer bestimmten Reihenfolge verschiedener »Begriffs-Impulse«.

Drei Hauptarten der »Begriffs-Impulse« ließen sich auf Grund umfangreicher Tonbandaufzeichnungen von Unterrichtsstunden in Sekundarschulen identifizieren: (1) Gesprächsteile, in denen Informationen über *Beispiele* des Begriffs besprochen wurden (»Beispiels-Impulse«) (instantial moves); (2) Gesprächsteile, in denen der Begriff mit anderen *verglichen* oder kontrastiert wurde (»Vergleichs-Impulse«) (comparative moves); und (3) die Gesprächsteile, in denen die Kriterien für den Begriff festgesetzt oder direkt *beschrieben* wurden (»Beschreibungs-Impulse«) (descriptive moves). Tabelle 1 enthält eine Liste dieser verschiedenen Arten der Impulse.

Tabelle 1
Impulsarten in »Begriffs-Abschnitten«*
(Types of Moves in Concept Ventures)

I Beschreibungs-Impulse 1. Beschreibung eines Merkmals 2. Definition ausreichender Bedingungen 3. Klassifikation des Begriffs 4. Klassifikatorische Beschreibung 5. Definition der Beziehungen zwischen Merkmalen 6. Zerlegung in Einzelteile II Vergleichs-Impulse 7. Vergleich durch Analogie 8. Vergleich durch Differenzierung 9. Vergleich von Beispielen eines Begriffs	III Beispiel-Impulse 10. Beschreibung positiver Beispiele 11. Aufzählung aller Beispiele 12. Beschreibung eines Nicht-Beispiels 13. Beschreibung der Produktion von Beispielen 14. Bestätigung einer Sache als Beispiel IV Gebrauchs-Impulse (usage moves) 15. Metaunterscheidungen bei Verwendung eines Begriffs
---	--

* aus: B. O. Smith, *The Strategies of Teaching* (1967)

Verschiedene Lehrer wurden hinsichtlich der in ihren Klassen üblichen Sequenzen der Impulse beschrieben und verglichen. Dabei stellte sich z. B. heraus, daß eine wechselnde Sequenz der Erörterung von Kriterien und die positiven Beispiele eines Begriffs eine der verbreitetsten Unterrichtsstrategien ist.

Unterschiede im Lehrerverhalten beim gleichen Curriculum

Die Praxis vieler Forscher, irgendwelche leicht zugänglichen Stichproben von Lehrern oder Klassen zu untersuchen, hat einige Unzufriedenheit geschaffen. Neuere Untersuchungen berichten dagegen von sorgfältig erhobenen Daten. Danach kann nun einiges darüber ausgesagt werden, wie das Verhalten im Unterricht durch gewisse, als unabhängig bekannte Variablen verändert werden kann. Gallagher (1966) berichtete über einen interessanten Versuch, die Unterschiede im Unterrichtsverhalten von Lehrern zu beschreiben, die dasselbe Thema desselben Curriculum mit vergleichbaren Klassen behandeln. Gallagher und andere (1966) entwickelten ein System der »Themenanalyse« (topic analysis), das die unterschiedlichen Gesprächsgänge über das gleiche Thema in verschiedenen Klassen identifizieren sollte. Als »Thema« wurde der Teil der Diskussion bezeichnet, in dem »sich die Klassendiskussion auf eine gegebene Handlung, einen Begriff

oder ein Prinzip konzentriert«. Jedes Thema wurde nach der Art des Inhalts klassifiziert (Informationswissen, Verfahrensfertigkeiten), nach dem Grad der Abstraktheit in der Diskussion (Daten, Begriff, Generalisierung) und nach dem Denkstil, der dem Gespräch zu entnehmen ist (Beschreibung, Expansion, Erklärung, Rechtfertigung, Bewertung usw.). Mit diesem System wurden die Tonbandprotokolle der Diskussionen in den Klassen von sechs Biologielehrern analysiert, die das Thema »Photosynthese« mit Hilfe der Materialien der neuen Biological Science Curriculum Study (»blue book version«) behandelt hatten. Gallagher fand signifikante Unterschiede zwischen den Lehrern bezüglich des Inhalts der Themen und im Grad der Abstraktheit, nicht aber im Denkstil. Auch der Anteil der Zeit, in der der Lehrer sprach, variierte signifikant von Klasse zu Klasse. Signifikante Unterschiede gab es auch für drei der vier Denkart: Beschreibung, Erklärung und Expansion. Gallaghers Analyse machte deutlich, daß auch bei konstant gehaltenem Curriculum eine große Vielfalt in den behandelten Themen und in der Art ihrer Behandlung auftritt.

*Unterschiede im Unterrichtsverhalten in bezug auf das Schuljahr,
das Alter des Lehrers und die Unterrichtsinhalte.*

Weitere Informationen über die Art der Unterschiede zwischen Lehrern liefert der Bericht einer Untersuchung von Biddle und Adams (1967). Diese Wissenschaftler machten Fernsehaufzeichnungen von 32 Unterrichtsstunden, die von 16 Lehrern in drei verschiedenen Schuljahren (im ersten, sechsten und elften) gehalten wurden. Diese Unterrichtsstunden wurden gleichmäßig auf Mathematik und Sozialkunde verteilt; ebenso bildete man aus den Lehrern zwei gleich große Gruppen (älter oder jünger als 30 Jahre). Die Zusammenfassung einiger ausgewählter Ergebnisse soll hier wiedergegeben werden:

(1) *Schuljahr.* Viele Unterschiede ergaben sich als eine Funktion des Schuljahrs. Die ersten Klassen verbrachten viel Zeit mit Handlungen (wie Singen, Vorlesen) und mit Anweisungen zur Organisation des Unterrichts. Die sechsten Klassen wirkten am wenigsten traditionell, da es in ihnen mehr Gruppenarbeit und viele Interaktionen gab. Erwartungsgemäß waren die elften Klassen stärker thematisch orientiert und verbrachten mehr Zeit für die intellektuelle Auseinandersetzung mit relevanten Unterrichtsinhalten.

(2) *Alter des Lehrers.* Die Klassen der älteren Lehrer schienen insofern traditioneller zu sein, als ihre Handlungen öfter vom Lehrer gesteuert wurden. Sie beschäftigten sich auch mehr damit, Informationen weiterzugeben, als zu intellektuellen Auseinandersetzungen (z. B. zu klärenden und bewer-

tenden Diskussionen u. ä.) hinzufügen. Die Klassen mit jüngeren Lehrern zeigten weniger Lehrer-Kontrolle; hier waren die Lehrer öfter mit kleinen Arbeitsgruppen beschäftigt oder überhaupt nicht mit einbezogen.

(3) *Fachspezifische Inhalte.* Die Diskussion in den Mathematikstunden neigte dazu, in den unteren Klassen einen geringeren und in der elften Klasse einen höheren »intellektuellen« Gehalt zu haben als in den entsprechenden Unterrichtsstunden der Sozialkunde. Bei den letzteren gab es dagegen mehr Diskussionen über irrelevante Themen.

Interessante Zusatzbefunde:

(4) Wenn der Lehrer als Sender (Redner) vor der ganzen Klasse spricht, besteht der Inhalt seiner Rede aller Wahrscheinlichkeit nach in der Vermittlung einfacher Informationen über das Thema oder aus Anweisungen zur Organisation des Unterrichts.

(5) Wenn der Lehrer etwas zum Thema vorträgt, steht er wahrscheinlich vor der Klasse; wenn er dagegen etwas zur Organisation des Unterrichts äußert, steht er oft seitlich zur Klasse.

(6) Die Schüler, die in der Klasse entlang der Mittellinie sitzen, nehmen eher an der Diskussion teil und werden häufiger vom Lehrer angesprochen. Die Schüler, die weiter von der Mittellinie weg sitzen, sprechen entschieden seltener und werden auch seltener angesprochen.

Diese kurze Zusammenfassung kann den vielen Daten, die Biddle und Adams sammelten, zwar kaum gerecht werden; sie soll jedoch auf die allgemeine Bedeutung dieser umfassenden Dokumentation struktureller und funktioneller Aspekte des Unterrichtsverhaltens hinweisen.

Verbale Lehrerreaktion

Deutlich unterscheidet sich von der Studie von Biddle und Adams eine Untersuchung, über die Zahorik (1968) berichtete. Während Biddle und Adams möglichst viele Aspekte des Unterrichtsverhaltens berücksichtigten, ohne ihrer Analyse dabei eine bestimmte Richtung zu geben, richtete Zahorik seine Aufmerksamkeit auf einen einzigen Aspekt, der zweifellos von beträchtlicher Bedeutung ist. Zahorik entwickelte ein System zur Analyse und Bewertung der Reaktionen des Lehrers auf Schülerantworten. Zahoriks Daten wurden aus Tonbandprotokollen gewonnen, die in acht Klassen des dritten und in sieben Klassen des sechsten Schuljahrs aufgezeichnet worden waren. Man hatte die Lehrer gebeten, den Inhalt eines aktuellen Nachrichtenmagazins zu behandeln, wodurch man eine gewisse Kontrolle über den Unterrichtsstoff in den beiden protokollierten Stunden erhielt. Die erste Stunde sollte eine Einführung in die Lektüre des Nachrichtenmagazins geben, und die zweite Stunde sollte dann daran anknüpfen.

Aus den Reaktionen der Lehrer auf die Antworten ihrer Schüler bildete Zahorik 14 verschiedene Kategorien. Sie werden in Tabelle 2 aufgeführt:

Tabelle 2
Kategorien der verbalen Lehrer-Reaktionen *

-
1. Lobende Bestätigung
 2. Tadelnde Verneinung
 3. Lobende Bestätigung + Tadelnde Verneinung
 4. Positive Antwort
 5. Negative Antwort
 6. Positive Antwort + Negative Antwort
 7. Positive Erklärung
 8. Negative Erklärung
 9. Erweiterung der Reaktion: Entwicklung
 10. Erweiterung der Reaktion: Verbesserung
 11. Wiederholung der Aufforderung: mehrere Antworten
 12. Wiederholung der Aufforderung: eine Antwort
 13. Fortsetzung des Unterrichts: neues Thema
 14. Verschiedenes
-

* aus: J. A. Zahorik, Classroom Feedback Behavior of Teachers, *Journal of Educational Research* 62, 1968, 147-150.

Die protokollierten Lehrerreaktionen bestanden aus einer oder einer Sequenz solcher Äußerungen. Außer den Angaben über die Häufigkeit dieser verschiedenen Reaktionen verschaffte sich Zahorik auch noch Lehrer-Einschätzungen (ratings) über die Richtigkeit der Schülerantworten und Schüler-Einschätzungen (ratings) über den Wert einer Auswahl von Lehrerreaktionen. Seine Ergebnisse wiesen darauf hin, daß von den 175 verschiedenen Arten von Lehrerreaktionen nur sechzehn häufiger wiederkehrten. Am häufigsten kamen folgende vor:

- (1) Positive Antwort mit anschließender Überleitung zu neuem Thema (8,5 Prozent)
 - (2) Bitte um Erweiterung der Antwort, ohne irgendwelche Hilfen zu geben (8,3 Prozent)
 - (3) Einfache lobende Bestätigung mit anschließender Überleitung zu neuem Thema (7,8 Prozent)
 - (4) Einfache lobende Bestätigung, dann positive Antwort (Wiederholung), dann Überleitung zu neuem Thema (5,8 Prozent)
 - (5) Keine Antwort, Überleitung zu neuem Thema (5,1 Prozent)
- Die Häufigkeit einer Reaktionsart variierte mit der Schulklasse, mit der

Phase des Unterrichts und mit den Urteilen der Lehrer über die Richtigkeit der Antwort. Durch diese Untersuchung konnte deutlich gemacht werden, daß die Rückmeldung durch den Lehrer keine Verstärkung (reinforcement) in dem Sinne ist, in dem Psychologen wie Skinner diesen Begriff verwenden. Auch betrachten die Lehrer Schülerantworten nicht als ein Verhalten, das Verstärkung braucht. Diese Untersuchung hat ein überraschendes Ergebnis: »Sie zeigt, daß es dem Lehrer nicht in erster Linie auf das sofortige Lernen ankommt und daß es ihm nicht das Wichtigste ist, was das Kind während der Interaktion sagt und tut« (Zahorik 1968).

Diese Ergebnisse sollten keinen Lehrer verwundern, es sei denn, er hätte sich von den Schriften einiger behavioristischer Theoretiker davon überzeugen lassen, daß Schüler nur das lernen, was sie tun, und daß ihr Tun aktiv verstärkt werden muß. Zur Verstärkung des Unterrichtsverhaltens sind viele vage Generalisierungen vorgeschlagen worden, besonders in Hinsicht auf das verbale Verhalten von Schülern. Zahoriks Untersuchung sollte dazu beitragen, entstandene Mißverständnisse zu beseitigen.

Damit soll freilich nicht gesagt werden, die Reaktionen des Lehrers auf die Antworten der Schüler seien unwichtig. Eine Untersuchung von Emmer (1968) weist darauf hin, daß einige Kategorien der Lehrerreaktionen (z. B. Flanders' Kategorie 3: »akzeptiert oder verwendet die Ideen des Schülers«) damit zusammenhängen, wie oft die Schüler durch Fragen die Initiative ergreifen. Emmers Ergebnisse, die aus der Untersuchung von sechzehn Lehrern des zweiten Schuljahres stammen, sind nicht unwidersprochen geblieben; sie machen jedoch deutlich, daß die Schüler um so freier mitmachen, je mehr der Lehrer ihre Ideen akzeptiert und darauf aufbaut. Wenn man davon ausgeht, daß es eine Beziehung zwischen der Beteiligung der Schüler an der Unterrichtsdiskussion und ihrer Leistung gibt – und einiges weist darauf hin (z. B. Gallagher 1966) –, dann sollte man auch nach der Beziehung zwischen der Art der verbalen Lehrerreaktionen und der Schülerleistung fragen. Das bedeutet jedoch nicht einfach, daß Schüler sich besser daran erinnern, was sie im Unterricht getan oder gesagt haben.

Unterschiede im Lehrerverhalten während des ersten Berufsjahrs

Medley berichtete über Beobachtungen an 70 Lehrern der Sekundarschule während des ersten halben Jahres ihrer Berufspraxis. Das verwendete Beobachtungssystem war das System »Observation Schedule and Record« (OScAR 4V), das Medley und seine Mitarbeiter im Laufe mehrerer Jahre entwickelt hatten. Als signifikante Veränderung im Verhalten dieser Lehrer ergab sich: (1) sie neigten dazu, weniger ergänzende Fragen (d. h., anders formulierte oder zusätzliche Fragen) zu stellen, (2) weniger Reaktionen auf

Schülerantworten zu geben, und (3) mehr Antworten als falsch abzulehnen, gleichzeitig jedoch die Schüler zum Antworten aufzumuntern. Diese Ergebnisse geben vielleicht die zunehmende Geschicklichkeit der Lehrer bei der Formulierung und bei der Auswahl des rechten Zeitpunkts ihrer Fragen wieder. Sie sehen immer seltener einen Anlaß, sich selbst zu verbessern, weil sie immer eindeutigere und angemessenere Fragen stellen.

Die Beziehung zwischen Lehrerverhalten und Schülerleistung

Affektive Dimensionen der Unterrichtsinteraktion

Das Unbehagen nach früheren Mißerfolgen und der Glaube daran, daß alle wichtigen Variablen im Unterricht noch entdeckt werden müssen, haben in neuerer Zeit Forscher dazu geführt, an das Problem der Ursachen von Schülerleistung nur mit beträchtlicher Vorsicht heranzugehen. Ohne Widerspruch zu fürchten, behauptete Bloom in einer Rede über den Stand der pädagogischen Forschung (1966), die Erforschung von Lehrmethoden habe gezeigt, daß die meisten etwa gleich wirkungsvoll sind. Dabei waren einige bedeutende Unterschiede bereits damals bekannt. Flanders (1965) hatte schon die Ergebnisse seiner umfassenden Untersuchung berichtet, in der er eine eindeutige Beziehung zwischen seinem Maß »Indirektheit« in der Unterrichtsinteraktion und der Schülerleistung in der Sozialkunde und in der Mathematik hatte zeigen können. Seinen Untersuchungen zufolge schien es so, als-ob die »indirekteren« Lehrer ein Unterrichtsklima schufen, in dem die Schüler Sympathien für ihre Lehrer empfanden und in standardisierten Leistungstests besser abschnitten.

In einem kleinen Kontrollversuch bestätigte La Shier (1967) diese Ergebnisse. La Shier gebrauchte das Flanderssche Verfahren zur Bewertung der Leistung von zehn Schulpraktikanten (student teachers), die Schüler der achten Klasse in einer sechswöchigen Arbeitseinheit über »tierisches Verhalten« aus dem Biological Science Curriculum Study (BSCS) unterrichteten. Den Schülern wurde in den entsprechenden Inhalten der Biologie ein Vor- und ein Nachtest gegeben, und sie wurden gebeten, einen Fragebogen über ihre Einstellungen zu dem Unterricht auszufüllen. Als Kriteriumswert diente die durchschnittliche Verbesserung der mittleren Leistung jeder Klasse im Vergleich zur Ausgangsleistung (California Test of Mental Maturity). Die Ergebnisse wiesen darauf hin, daß die »Indirektheit« (I/D Verhältnis) der Praktikanten in signifikanter Beziehung zum Leistungsgewinn und den positiven Einstellungen bei den Schülern stand.

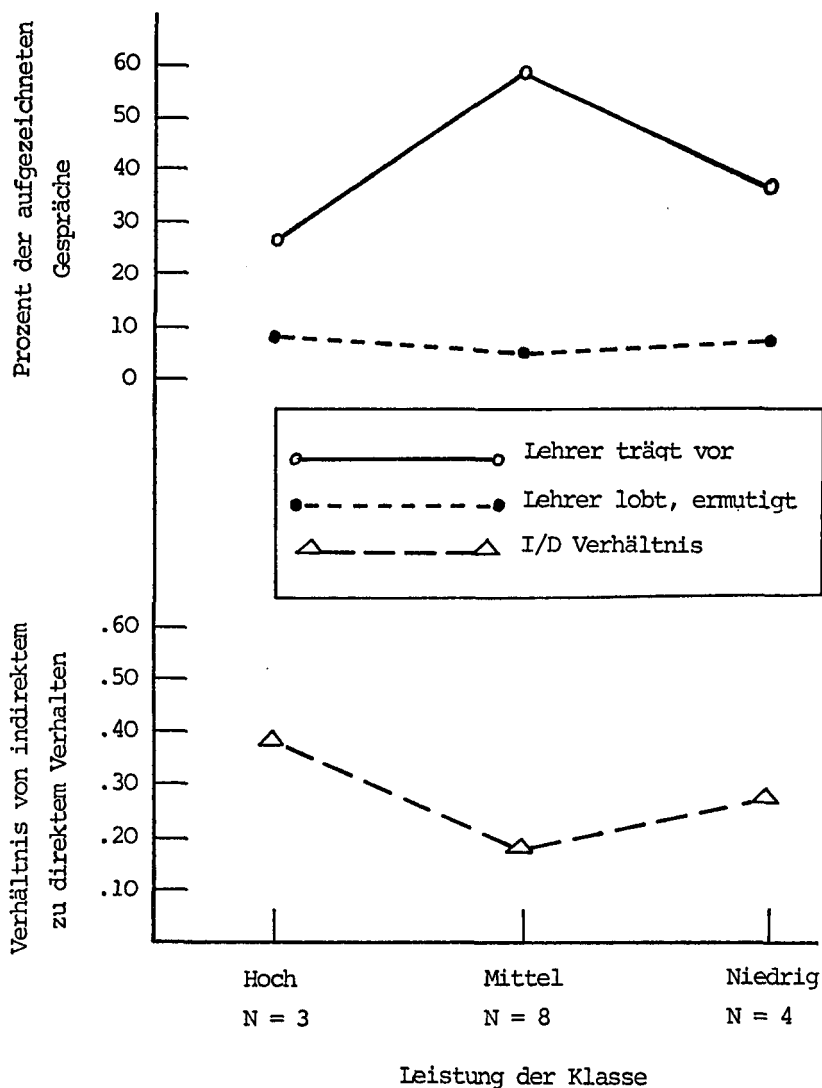
Das Verhältnis zwischen der Einstellung der Schüler und ihrem Leistungszuwachs war ebenfalls signifikant.

In einer noch umfassenderen Untersuchung benutzte Furst (1967) Daten von Bellack und anderen (Bellack u. a. 1966) in dem Bemühen, die Schülerleistung mit Maßen des Flandersschen Verfahrens der Interaktionsanalyse und mit Maßen des Bellackschen Analyse-Systems in Beziehung zu bringen. Bellacks Daten bestanden aus schriftlichen Protokollen von je vier Unterrichtsstunden, die fünfzehn Lehrer in ihren New Yorker Sekundarschulen gehalten hatten. Der Unterrichtsstoff war ein standardisierter vierstündiger Kurs in Wirtschaftswissenschaften. Auf Grund der Ergebnisse der Leistungstests (korrigiert nach Schülerintelligenz und Klassengröße) wurden die Lehrer aufgeteilt. Drei Lehrer wurden den Klassen mit den besten Leistungen, vier denen mit den schwächsten Leistungen und acht denen mit den mittleren Leistungen zugeordnet. Diese drei Lehrergruppen wurden nach drei Maßen des Flandersschen Verfahrens und nach drei des Bellackschen Analyse-Systems miteinander verglichen. Die Ergebnisse zeigten, daß die drei Klassen mit der höchsten Leistung sich von den anderen in folgendem unterschieden: mehr ausführliche, indirekte Gesprächsbeiträge der Lehrer, mehr positive als negative unmittelbare Reaktionen auf Schülerantworten, intensivere Unterrichtsbeteiligung seitens der Schüler. Bei näherer Betrachtung zeigen Fursts Ergebnisse freilich keine einfache Beziehung zwischen diesen Maßen und der Schülerleistung. Eine Auswahl der Ergebnisse wird in Abbildung 1 dargestellt. Hieraus ist zu ersehen, daß eine Kurvilinearität zwischen der Leistung und dem Prozentsatz der verschiedenen Arten von protokolliertem Lehrer- und Schülerverhalten besteht. Ob diese deutliche Kurvilinearität in Fursts Daten allerdings einmalig ist oder ob dieser Sachverhalt von anderen Forschern nur übersehen worden ist, ist nicht bekannt. Aber das weist darauf hin, daß bei einer anderen Lehrer-Stichprobe möglicherweise ganz andere Ergebnisse gefunden worden wären.

Noch einige Ergebnisse von Furst müssen erwähnt werden. Sie legten die Hypothese nahe, Lehrer mit hohen Leistungsergebnissen gäben in ihren Unterrichtsdiskussionen eine mittlere Anzahl pädagogischer Impulse zur verbalen Strukturierung des Unterrichts (verbal structuring moves), zeigten ein mittleres Verhältnis im Frage-Antwort-Austausch und einen hohen Grad an Verschiedenheit in den logischen Prozessen. Ein zusammengesetzter Meßwert, der aus den Einzelwerten dieser »kognitiven« Variablen bestand, bestätigte diese Hypothese. Ohne weitere Details über die Art dieser Meßwerte und ihre Beziehung zur Leistung ist es jedoch unmöglich, die Bedeutung dieser Ergebnisse richtig einzuschätzen.

Powell (1968) versuchte zu zeigen, daß man auf Grund der Beobachtun-

Abbildung 1. Meßwerte der Klassen mit hohen, mittleren und niedrigen Leistungen, gewonnen mit dem Flandersschen Verfahren der Interaktionsanalyse (aus: Furst, 1967).



gen von Lehrern in ihrer normalen schulischen Umgebung auf signifikante Beziehungen zwischen »Indirektheit« und besseren Schülerleistungen bei standardisierten Leistungsprüfungen schließen kann. Er verwendete in verschiedenen Klassen des dritten Schuljahres das Flanderssche Verfahren und wiederholte das Ganze im folgenden Jahr in denselben Klassen (viertes Schuljahr). Er wählte die Schüler aus, die während ihrer ersten drei Schuljahre dieselben »indirekten« bzw. »direkten« Lehrer hatten, und im vierten Jahr dann umgekehrt bei »direkten« bzw. »indirekten« Lehrern waren. Die Ergebnisse zeigen, daß die Leistung im Rechnen, nicht aber im Lesen, mit der »Indirektheit« des Lehrers während der ersten drei Schuljahre eng zusammenhängt. Und im vierten Schuljahr hatte es offenbar überhaupt keine Wirkung, ob die Schüler bei einem »direkten« oder »indirekten« Lehrer waren. Der Autor schloß daraus:

Von den Ergebnissen dieser Untersuchung her scheint es eindeutig, daß der indirekte Unterricht keinen deutlichen Gesamtvorteil bringt (Powell 1968, 4).

Das Problem bei solchen Untersuchungen liegt in der Neigung des Forschers, eine Eigenschaft wie »Indirektheit« als feststehendes Attribut zu betrachten. Wenn der eine Forscher die beobachteten Lehrer in die »indirektesten« und die »direktesten« einteilt, kann seine Trennungslinie an einer anderen Stelle liegen als die eines anderen Forschers, der eine andere Lehrerstichprobe untersucht. Die Lehrer, die in der einen Untersuchung als »indirekt« bezeichnet werden, können in einer anderen Untersuchung als »direkt« bezeichnet werden. Dies macht vielleicht nicht sehr viel aus, sofern das Verhältnis zwischen den Faktoren des Unterrichtsklimas und der Leistung einfach linear ist. Aber wenn dieses Verhältnis kurvilinear ist, wie die Daten in Fursts Untersuchung andeuten, dann wird eine sorgfältigere Analyse der Daten erforderlich.

Der vielleicht sorgfältigste und umfangreichste Versuch, Aspekte des Unterrichtsklimas und des Lehrereinflusses mit der Schülerleistung in Beziehung zu bringen, wurde von Soar (1967; 1968) berichtet. Soar begann seine Untersuchung mit der Sammlung umfangreicher Daten über die Schüler in 55 Klassen (drittes bis sechstes Schuljahr) in vier verschiedenen Elementarschulen. Er ergänzte die standardisierten Leistungstests für Wortschatz, Lesen und Rechnen mit Kreativitätstests (Minnesota), Angsttests (Children's Manifest Anxiety Scale), Tests über Neigung zur Abhängigkeit (Flanders 1960) und Fragebogen, mit deren Hilfe die Einstellungen der Schüler gegenüber Lehrern und Mitschülern abgeschätzt werden sollten. Die gleichen Tests wurden nach einem Jahr und am Ende des zweiten Jahres wiederholt. Die Werte des Lernzuwachses der Schüler über den Zeitraum von einem bzw. zwei Jahren wurden als Kriteriumswerte verwendet.

In jeder der 55 Klassen wurden Beobachtungen angestellt, und zwar mit dem Flandersschen Verfahren der Interaktionsanalyse und mit einem anderen Verfahren, das Teile von Medley und Mitzels OSCAR und Fowlers Ablehnungs-Zustimmungs-Skala (Hostility Affection Schedule) einschloß. Alle Meßwerte aus diesen Beobachtungen wurden faktorenanalysiert; es wurde eine Neun-Faktoren-Lösung erarbeitet, die »das Bild vom Unterricht am deutlichsten wiedergab«.

Für jede Klasse wurde ein Faktorenwert für jeden dieser neun Faktoren errechnet. Diese Faktorenwerte wurden dann mit dem durchschnittlichen Lernzuwachs korreliert, den die Schüler in den Tests erreicht hatten. Die Ergebnisse können so zusammengefaßt werden:

(1) Es gab keinen signifikanten Zusammenhang zwischen den beiden Faktoren, die die Ursache für die höchsten Prozentwerte der Varianz waren, und einem der Maße des Leistungszuwachses. Diese Faktoren schienen vielmehr Maße für die Gesamtdauer der sprachlichen Äußerungen der Lehrer und für die Möglichkeit der Schüler zu sein, eine Diskussion anzufangen.

(2) Der Faktor mit der engsten Beziehung zur Leistungsverbesserung der Schüler wurde als »ausgedehnter Lehrervortrag im Gegensatz zu schneller Lehrer-Schüler-Interaktion« bezeichnet (extended discourse vs. rapid teacher-pupil interchange). Soar bemerkt dazu:

»Hier spiegelt sich offenbar ein Zyklus wider, in dem zuerst der Lehrer 15 bis 20 Sekunden lang redet, eine Frage stellt oder eine Anweisung gibt, und dann die Schüler eine Weile lang reden, . . . kein längerer Vortrag im üblichen Sinn also, sondern eine Reihe von Lehreräußerungen, die der Problem-Strukturierung dienen können, oder die Darbietung einer begrenzten Zahl von Informationen, die die Schüler weiterentwickeln und verwenden sollen.« (Soar 1967, 7).

Die Parallele zwischen dieser Beschreibung und Fursts »mittlerem Grad der Steuerung der Lernaktivitäten« ist evident.

(3) Die Häufigkeit verbaler Schärfe oder Kritik durch den Lehrer stand in negativer Beziehung zum Lernzuwachs der Schüler im Rechnen und in einiger Beziehung zu ihrer Angst und ihrer Neigung zur Abhängigkeit.

(4) Ein Faktor zeigte viel Ähnlichkeit mit Flanders Maß der »Indirektheit«. Er stand jedoch weniger zur »Direktheit« des Lehrers in negativer Beziehung als vielmehr zum Ausmaß an »Ruhe und Unruhe« im Unterricht. Als Faktor korrelierte er zwar nicht mit der Verbesserung der Schülerleistung, stand aber in positiver Beziehung zur Steigerung der Kreativitätswerte und der Interessen der Schüler am Unterricht. Soar fiel es schwer zu erklären, warum der Faktor, der am engsten mit der »Indirektheit« des Lehrers zusammenhing, keine Korrelation mit der Leistung zeigte. Er

korrelierte jedoch mit dem jeweiligen Schuljahr, wobei Soar davon ausging, daß bei einigen früheren Untersuchungen man vielleicht das Ausmaß der Leistung mit der Höhe des Schuljahrs vermischt habe.

Im zweiten Teil seiner Untersuchung erforschte Soar den Einfluß des Lehrerverhaltens auf die Entwicklung der Schüler im Laufe eines zweiten Jahres. Er interessierte sich für das Ausmaß, in dem sich die Wirkungen des Lehrerverhaltens eines bestimmten Jahres vielleicht auf nachfolgende Jahre übertragen. Die Ergebnisse zeigten, daß die Kritik der Lehrer nur in geringem Maße von einem Jahr auf das nächste übertragen wird:

Es scheint, als ob Ausdrücke negativer Affekte im Unterricht ihren stärksten Einfluß in der ersten Zeit danach hätten. Diesjährige Kritik ist von Bedeutung, die letztjährige nicht mehr.

Die Ergebnisse für die anderen Faktoren, die die Schülerleistung beeinflussen, zeigen kontinuierliche Auswirkungen und eine gewisse Interaktion zwischen den beiden Jahren. Der wohl auffallendste Befund im Hinblick auf die oben berichteten Daten von Furst besteht darin, daß ein mittleres Ausmaß von Lehrerkontrolle und »Indirektheit« die größten Vorteile zu bringen schien:

... in diesen Klassen erzeugte ein mittleres Maß an Kontrolle, entweder in Form von Kritik oder indirektem Unterricht, mehr wünschenswerte Veränderungen im Schülerverhalten als der übertriebene Mangel an Lehrerkontrolle. Vielleicht läßt sich daraus die Notwendigkeit ersehen, daß der Lehrer ein Minimum an Strukturierung schaffen muß, innerhalb derer die Schülerleistung maximiert werden kann (Soar 1967, 10).

Zweifellos wird es weitere Untersuchungen geben, die den Versuch machen werden, Schülerleistung und Wert für das affektive Klima im Unterricht miteinander in Beziehung zu bringen. Aber schon jetzt scheint genügend Beweismaterial vorzuliegen, um einige allgemeine Schlüsse ziehen zu können.

Erstens: Beobachtungsverfahren mit nur einem einzigen Kriterium wie dem Ausmaß der »Indirektheit« des Lehrers sind wahrscheinlich in ihrem Ansatz zu ungenau, um jemals klare Beziehungen mit Kriterien der Schülerleistung aufzeigen zu können. In einer Reihe von Beobachtungen werden sie wahrscheinlich zufällig mit mehreren anderen Aspekten des Unterrichtsverhaltens in Verbindung gesetzt.

Zweitens: Selbst wenn diese Beobachtungsverfahren eine gültige Dimension des Unterrichtsverhaltens erfassen, ist es unwahrscheinlich, daß diese Dimension eine einfache lineare Beziehung zur Schülerleistung aufweist. Wahrscheinlich ist es naiv zu erwarten, daß eine größere Freiheit der Schüler und ein freundlicheres Verhalten der Lehrer direkte und kontinuierliche Fortschritte in der intellektuellen Entwicklung des Schülers bringen.

Wie aus Soars Untersuchung deutlich hervorgeht, ist Kritik und verbale Schärfe der Lehrer gegenüber den Schülern ein wichtiger Faktor, aber das Ausbleiben dieser Kritik führt nicht automatisch zur Verbesserung des Lernens. Untersuchungen über das Unterrichtsklima sind schließlich Untersuchungen der Bedingungen, *unter denen Lernen stattfindet*. Daß es eine beträchtliche Vielfalt an Bedingungen gibt, die das Lernen nicht verhindern, sollte dabei nicht überraschen.

Die Suche nach signifikanten kognitiven Variablen

In einer neueren Abhandlung über die kognitiven Aspekte im Unterricht, bezweifelte Gage (1966 b) die Nützlichkeit, solche deskriptiven Verfahren wie die Systeme von Bellack und Smith zu entwickeln, bevor etwas über die mögliche Gültigkeit der kognitiven Aspekte des Unterrichtens bekannt ist, die analysiert werden sollen. Er zitierte eine Untersuchung, in der offensichtlich signifikante Unterschiede zwischen den Fähigkeiten der Lehrer, den Schülern ein Prinzip zu erklären, auftreten. Gage schlug einen systematischen Vergleich des Verhaltens solcher Lehrer vor, die bekannterweise gut bzw. schlecht erklären können; dieser Vergleich sollte wichtige Hinweise auf die wesentlichen Elemente solchen Unterrichts liefern. Im Anschluß daran berichteten Fortune, Gage und Shutes (1966) über die Ergebnisse eines Vergleichs der »Erklärungsfähigkeit« von 40 Praktikanten. Diese Praktikanten mußten in viertelstündigen Lektionen kleinen Gruppen von Schülern der Sekundarschule eine kontrollierte Themenreihe unterrichten. Als Kriteriumswert für jedes Thema wurde ein Test mit zehn Auswahl-Antwort-Aufgaben benutzt (multiple-choice item test), von denen die Praktikanten die Hälfte schon gesehen hatten, die andere Hälfte ihnen aber unbekannt war. Die Untersuchung war so aufgebaut, daß die Fähigkeit der Praktikanten zu erklären bzw. die Themen zu unterrichten über die verschiedenen Themen und die verschiedenen Schülergruppen hinweg verglichen werden konnte. Die Ergebnisse wiesen darauf hin, daß sich das Erklärungsvermögen eines Lehrers wahrscheinlich mit verschiedenen Themen verändert, bei verschiedenen Schülergruppen jedoch relativ konstant bleibt.

Vor einiger Zeit machten Gage und seine Mitarbeiter Fernsehaufzeichnungen von 43 Lehrern der Sozialkunde, die Schülern der zwölften Klasse Unterricht in zwei standardisierten Themen gaben. Die Themen bezogen sich auf aktuelle ökonomische, politische und soziale Entwicklungen in Jugoslawien und Thailand. Die Lehrer unterrichteten beide Themen in Lektionen von 15 Minuten, während ein drittes Thema (über Israel) allen Klassen in Form einer standardisierten Tonbandlektion vorgeführt wurde. Die Fähigkeit der Schüler, den Inhalt der standardisierten Tonbandlektion zu

lernen, diente dazu, die Kriteriumstestwerte an die anderen zwei Themen anzupassen. Die Schüler in jeder Klasse wurden außerdem darum gebeten, anhand standardisierter Fragebogen jede Lektion zu bewerten und den Grad ihrer Aufmerksamkeit während jeder Lektion anzugeben.

Ist die Erklärungsfähigkeit der Lehrer für verschiedene Themen relativ konstant? Podlogar, Rosenshine und Gage entdeckten (1968) in ihrer Datenanalyse, daß die Korrelation zwischen den durchschnittlichen Testwerten der Schüler für beide Themen zwischen .41 und .47 lag. Diese Korrelation ist signifikant und weist darauf hin, daß zwischen 16 und 20 Prozent der Varianz in der Klassenleistung auf einen Faktor der Fähigkeit des Lehrers zurückzuführen sein könnte. Aus den Bewertungen der Schüler wurde außerdem klar, daß sie die Erklärungsfähigkeit eines Lehrers ziemlich genau einschätzen konnten und daß die Einstufung ihrer eigenen Aufmerksamkeit signifikant mit dem Ausmaß ihres Lernens korrelierte.

Zwei etwas verschiedene Versuche, die wichtigen Komponenten der »Erklärungsfähigkeit« zu isolieren, sind mit diesen Daten gemacht worden. Hiller, Fisher und Kaess (1968) sowie Dell und Hiller (1968) berichteten von einem Versuch, mit Hilfe eines Computers die kritischen Elemente im verbalen Verhalten von Lehrern zu isolieren. Rosenshine (1968) berichtete die Ergebnisse eines Vergleichs innerhalb einer Teilstichprobe, die aus den erfolgreichsten und erfolglosesten Lehrern der Gesamtstichprobe gebildet worden war; dabei wurde sowohl das Verbal- als auch das Handlungsverhalten der Lehrer untersucht. Die Computer-Analyse erbrachte zwei Dimensionen des verbalen Lehrerverhaltens, die verläßlich mit dem Lehrererfolg bei den beiden Lektionen zusammenzuhängen schienen. Die erste dieser Dimensionen wurde als »verbale Flüssigkeit« (verbal fluency) bezeichnet. Sie bestand in einem zusammengesetzten Maß, das auf der Durchschnittslänge des gesprochenen Satzes und auf anderen Kennzeichen der Sprachflüssigkeit, wie etwa dem Anteil der »ähs« in der Rede des Lehrers basiert. Der zweite Faktor wurde »Unbestimmtheit« (vagueness) genannt: gleichfalls ein zusammengesetztes Maß, das sich aus der relativen Häufigkeit solcher Worte wie »fast, im allgemeinen, vielleicht usw.« errechnet. Die Forscher folgerten daraus nicht, daß diese verbalen Faktoren in direktem Zusammenhang mit dem Lernen des Schülers standen:

Wir sind der Meinung, daß das Verhältnis zwischen unserem Maß und den Werten des Kriteriumstests vor allem die Korrelation und nicht die Ursache widerspiegelt. ... unser Maß der Unbestimmtheit dient wohl eher als Hinweis auf andere Verhaltensweisen, die ihrerseits das Verstehen und Behalten der Lektion kausal beeinflussen (Hiller/Fisher/Kaess 1968, 7).

In Rosenshines Analyse wurden viele Variablen untersucht; auf einige

ist bereits in früheren Untersuchungen hingewiesen worden, andere, z. B. die linguistischen Indikatoren, wurden erstmals hier untersucht. Die folgenden drei Variablen zeigten eine signifikante Beziehung zum Lehrererfolg:

(1) *Gesten und Bewegung*. Die fähigeren Lehrer zeigten eine größere Tendenz, sich frei in der Klasse zu bewegen und dabei Gesten zu machen.

(2) *Regel und Beispiel*. Die fähigeren Lehrer neigten dazu, Regeln vor und nach der Diskussion der Beispiele zu erklären, während die weniger fähigen Lehrer dazu neigten, die Regel nur einmal, entweder vor oder nach den Beispielen zu erklären.

(3) *Erklärende Bindeglieder* (explaining links). Die fähigeren Lehrer benutzten in der Regel häufiger Bindewörter wie »weil, weshalb, um . . . zu, folglich, mittels, usw.«.

Rosenshine beendete seinen Bericht mit der Behauptung, die wichtigsten Variablen seien offensichtlich jene, die mit der Struktur der Kommunikation des Lehrers zu tun hätten. Die sequenzielle Anordnung der Gedanken und ihrer tragenden Elemente bedarf weiterer Erforschung. Man sollte jedoch nicht zu viel in die Ergebnisse dieser von Gage an der Stanford Universität angefangenen Untersuchungsreihe hineinlesen. Es darf nicht vergessen werden, daß, obwohl die Daten auf der Untersuchung einer größeren Zahl von Lehrern beruhen, nur zwei fünfzehnminütige Stichproben von dem Unterricht eines jeden Lehrers gemacht wurden⁴. Die Ergebnisse dieser Analyse lassen sich jedoch ohne weiteres mit jenen anderer Forscher wie beispielsweise Furst (1967) und Soar (1966; 1967) vereinbaren.

Ein Versuch mit Variationen in der sequenziellen Anordnung von Unterrichtsinhalten

Eine Untersuchung, die aus der Arbeit von Smith und Meux (1967) entstand, wirft etwas Licht auf die Bedeutung der Gedankensequenz des Lehrers. In dieser Untersuchung (Nuthall 1968) wurden vier alternative Strategien zum Unterrichten von Begriffen (concept teaching strategies) bzw. Sequenzen von »Begriffs-Impulsen« (sequence of conceptual moves) miteinander verglichen (vgl. die Beschreibung von Smiths Analyse von Begriffs-Impulsen). Es waren vier Unterrichtsstrategien, die Smith und seine Mitarbeiter in ihren Tonbandaufzeichnungen von Unterrichtsstunden in Sekundarschulen identifiziert hatten. Der verbale Inhalt dieser Strategien wurde in programmierter Textform einer Stichprobe von 412 Sekundarschülern dargeboten. Eine Varianzanalyse der Ergebnisse nachträglicher Kriteriumstests (delayed criterion tests) deutete an, daß die alterna-

tiven Strategien signifikant andere Wirkungen hatten. Die weitere Analyse ließ vermuten, daß die Wirksamkeit einer Unterrichtsstrategie davon abhängig ist, in welchem Maße sie auf dem schon vorhandenen Wissen der Schüler aufbaut und wie sehr sich Unterrichtsstrategie und vorhandenes Wissen gegenseitig beeinflussen. Bei den beiden Begriffen, die während dieses Versuchs unterrichtet wurden, war der Gebrauch von Beispielen und von Vergleichen mit anderen Begriffen je nach dem Vorwissen der Schüler verschieden.

Einige Schlußfolgerungen und Hinweise für künftige Forschung

Wie zu Beginn dieses Beitrags erwähnt wurde, haben viele Untersuchungen der letzten Jahre Versuchscharakter gehabt. Man sollte nicht versuchen, endgültige Schlußfolgerungen zu ziehen, auf denen Lehrer ihre Praxis aufbauen oder woraus Forscher experimentelle Hypothesen ableiten könnten. Einige Ergebnisse scheinen jedoch in den Schlußfolgerungen vieler Forscher wiederzukehren. Alle beziehen sich auf verbales Lehrerverhalten als einer signifikanten Ursache für Schülerleistung. Solange der Lehrer den Unterricht von Feindseligkeit und übermäßiger Kritik freihalten kann, wird wahrscheinlich die Wirksamkeit seines Einflusses auf die Schülerleistung von solchen Dingen abhängig sein wie:

- (1) seiner Fähigkeit, den verbalen Kontext, in dem die Interaktion zwischen Lehrer und Schülern stattfindet, vorzubereiten bzw. zu »strukturieren«;
- (2) seiner Fähigkeit, bei der Darbietung der Unterrichtsinhalte Gedanken mit einem Maximum an logischem Zusammenhang und einem Minimum an Unbestimmtheit und Ziellosigkeit zu organisieren;
- (3) seiner Fähigkeit, die Schüler zur Teilnahme an Diskussionen anzuregen und sie für die Entwicklung und Erweiterung von Ideen zu interessieren.

Diese Themen sollten als Anregungen zu weiterer Forschung betrachtet werden, insbesondere für die Entwicklung experimentell brauchbarer Theorien des Unterrichtsverhaltens.

Dem kritischen Leser muß es aufgefallen sein, daß vielen der erwähnten Untersuchungen eine feste Richtung und eine kontrollierte Anordnung fehlen; dieser Mangel kann nur durch eine angemessene Theorie beseitigt

werden. Es drängt sich folgende Frage auf: Wenn diese Untersuchungen das Lernen der Schüler im Unterricht erfassen sollen, warum werden dann so wenige von ihnen durch die bekannten Lerntheorien beeinflusst? Wann wird eine Verbindung zwischen den etablierten psychologischen Theorien und dieser neuen Richtung der Erforschung des Unterrichtsverhaltens hergestellt?

Nach Meinung dieses Autors gibt es nur eine Antwort auf diese Fragen, daß nämlich die traditionelle psychologische Theorie keinen großen Wert haben kann, bevor die Unterrichtsforscher ihrerseits nicht signifikante theoretische Erklärungen des Unterrichtsgeschehens gefunden haben. Man braucht keine weitere Anpassung und Ausweitung bekannter Theorien, sondern die Schaffung einer neuen Theorie, die direkt auf das wirkliche Verhalten, das sie erklären soll, bezogen ist. Die in diesem Beitrag behandelten Untersuchungen weisen darauf hin, daß wir jetzt schon genug über einige wahrscheinlich wichtige Variablen wissen, um ihre Funktion wenigstens teilweise zu erklären.

Im Rahmen einer Unterrichtstheorie verdienen folgende Elemente des Unterrichtsverhaltens besondere Aufmerksamkeit:

(1) Nicht alle Schüler beteiligen sich immer am Unterricht; dennoch wird erwartet, daß alle lernen, und im allgemeinen lernen alle etwas. Eine Theorie des stellvertretenden Lernens (vicarious learning) beziehungsweise des Lernens bei nachlassender Teilnahme ist daher erforderlich.

(2) Die Bedeutung der Schülerbeteiligung sollte unabhängig von Begriffen, Reaktion (response) und Verstärkung (reinforcement) eingeschätzt werden. Viele Autoren (z. B. Smith 1961) haben angedeutet, daß Lehrer die Beteiligung ihrer Schüler an Diskussionen als Informationsquelle über den Wissensstand und die intellektuellen Prozesse der Schüler benutzen. Das heißt, sie diagnostizieren die Schülerreaktionen und treffen spontane Entscheidungen aufgrund dieser Diagnosen. Jacksons neuere Veröffentlichung (1968), die auf Interviews mit erfolgreichen Lehrern beruht, weist darauf hin, daß Lehrer ihre sichersten Erfolgszeichen der subtilen »Rückmeldung« entnehmen, die sie von den Schülern ihrer Klasse erhalten. Als Praktiker sind sie sehr mißtrauisch und voller Zweifel gegenüber den Ergebnissen standardisierter Tests. Soweit diesem Autor bekannt ist, hat bisher noch keiner die Vorstellungen vom Lernen zu erklären versucht, aufgrund derer ein fähiger Lehrer die Diskussion beeinflusst. Wie interpretiert er die Reaktionen seiner Schüler? Welche Zeichen benutzt er, um den Verlauf seiner Handlungen zu bestimmen? In den Antworten auf diese Fragen muß die Erklärung für Ursache und Wirkung der Diskussion im Unterricht liegen.

(3) Die Bedeutung des Schülers als einer unabhängigen Größe für Verän-

derungen im Unterricht muß berücksichtigt werden. In einem neueren Artikel hat Turner (1968) Material aufgearbeitet, welches die Hypothese unterstützt, daß Schüler zumindest einige Aspekte des Lehrerverhaltens selbst bestimmen. Schüler können ihrerseits ziemlich genau den Erfolg ihrer Lehrer bewerten (vgl. Podlogar/Rosenshine/Gage 1968); Beobachter des Unterrichtsverhaltens werden bald darauf aufmerksam, daß Schüler die subtilen Hinweise ihrer Lehrer begreifen können und es auch tatsächlich tun. Jede Theorie, die darauf zielt, das Lernen im Unterricht zu erklären, sollte die Schüler auch als aktiv auswählende Teilnehmer berücksichtigen.

(4) Die Doppelrolle des Lehrers, die darin besteht, z. T. intellektuell und psychisch auf die Schüler Einfluß zu nehmen (Gage, 1966a) und z. T. ihre Motivation zu erhöhen, bedarf weiterer Untersuchungen. Die aktive Teilnahme des Schülers zu erreichen ist eine oft genau so kunstvolle und schwierige Aufgabe wie die Führung der Gedankengänge von Schülern. Das Problem der Motivation im Unterricht muß bei jeder Untersuchung der gedanklichen Interaktionen mitberücksichtigt werden. Hier ist die Erzeugung von verdeckten geistigen Reaktionen (covert mental responses) ähnlich schwierig wie die Stimulierung offener verbaler Reaktionen.

Schließlich ist kein kritischer Bericht über die Unterrichtsforschung ohne die Warnung vollständig, die auch verschiedene andere Autoren geäußert haben: Die Frage nach den besten Unterrichtsmethoden und den wirksamsten Mitteln, das Lernen zu fördern, darf nicht mit den Fragen darüber gekoppelt werden, wie sich Unterrichten und Lernen im Unterricht in Wirklichkeit abspielen. Erst wenn wir über die Beziehungen zwischen Unterrichtsverhalten, Lernen und Leistung der Schüler genauere Kenntnisse haben, wird es vielleicht möglich sein, wohlbegründete Unterrichtsprinzipien aus diesem Verständnis abzuleiten. Dieser kritische Bericht sollte deutlich machen, daß ein solches Verständnis noch nicht erreicht worden ist. Wahrscheinlich werden die Bemühungen um ein ausreichendes Verständnis des Unterrichts zusätzlich erschwert, wenn dieser Wunsch nach einem besseren Verständnis mit dem nach einem verbesserten Unterricht verwechselt wird.

Abgedruckt auch in: M. C. Wittrock/D. E. Wiley, *The evaluation of instruction, issues and problems*, New York: Holt, Rinehart and Winston 1970, 221–238.

1 Sehr zum Leidwesen vieler unwilliger Beamter der Schulverwaltung gestehen wir jedoch daß dies ein tunlicher Ausgangspunkt wäre.

2 Manches spricht dafür, daß dies ein vernünftiges Verfahren ist. Vgl. J. S. Bekker 1962; Miller: *Income and Higher Education*, in: S. J. Muskin 1962; T. Schultz 1961.

GENE V. GLASS: Die Entwicklung einer Methodologie der Evaluation

Übersetzung von Hannes Graudenz (Dipl.-Psych.) und dem Herausgeber.

Originaltitel: *The growth of evaluation methodology*

Der deutschen Übersetzung liegt das dem Herausgeber vom Autor zugesandte Manuskript zugrunde.

1 Ein »allgemeines Phänomen« ist nachgewiesen oder kann entdeckt werden in einem weiten Feld von scheinbar verschiedenen Erscheinungen und wird als Kriterium zur Prüfung eines wissenschaftlichen Begriffs herangezogen. Ohne eine solche Qualifikation würde es bereits »Einschätzung wissenschaftlicher Wahrheit« bedeuten, empirisch festzustellen, daß man seine Schlüssel verloren hat. Der Begriff der Generalisierbarkeit von erwarteten Ergebnissen ist wichtig für die Unterscheidung von Evaluation und Forschung; er ist auch von großer praktischer Bedeutung beim Entwurf einer Evaluations-Untersuchung (vgl. Stake 1969).

2 In diesem Abschnitt beziehe ich mich weitgehend auf die Geschichte der North Central Association von Calvin O. Davis (1945).

3 Die Pionierarbeit von Joseph M. Rice mag manchem zu dieser Zeit bekannt gewesen sein, wurde aber wahrscheinlich eher als tendenzieller Journalismus denn als pädagogische Forschung angesehen.

ARNO A. BELLACK:

Methoden zur Beobachtung des Unterrichtsverhaltens von Lehrern und Schülern

Übersetzung von Dorothea Szymanski (Dipl.-Psych.).

Originaltitel: *Methods for observing classroom behaviour of teachers and students*, in: K. Ingenkamp: *Methods for the evaluation of comprehensive schools*, Weinheim, Berlin, Basel 1969, 187–215.

1* Zum gegenwärtigen Zeitpunkt liegen bereits 16 Bände mit Instrumenten zur Unterrichtsbeobachtung (79 Instrumente) vor (vgl. Bibliographie).

2* Vgl. dazu auch Wulf 1972a.

GRAHAM A. NUTHALL:

Ausgewählte neue Untersuchungen zur Unterrichtsinteraktion und zum Lehrverhalten: Ein kritischer Bericht

Übersetzung von J. Hermann (M. A.), G. Hermann (Dipl.-Psych.) und dem Herausgeber.

Originaltitel: A review of some selected recent studies of classroom interaction and teaching behavior.

American Educational Research Association, Monograph Series on Curriculum Evaluation, No. 6, Chicago: Rand McNally 1970, 6-29.

Dieser Beitrag erschien zuerst unter dem Titel: Types of research on teaching im New Zealand Journal of Educational Studies 3, 1968, 2.

1 Der Verfasser möchte sich bei den Professoren Nathaniel Gage (Stanford University) und Barak Rosenshine (University of Illinois, Urbana-Campaign), für verschiedene hilfreiche Hinweise bedanken.

2* Zum gegenwärtigen Zeitpunkt liegen bereits 16 Bände mit 79 Instrumenten zur Unterrichtsbeobachtung vor (vgl. Bibliographie).

3 Persönliche Mitteilung

4 Als dieser Artikel bereits geschrieben war, hat Professor Gage mich darauf hingewiesen, daß die Frage der zeitlichen Dauer der Stichprobe *irrelevant* sei, »insofern als das Kriterium der Effektivität tatsächlich unbeeinflusst ist oder entsprechend den Schülereigenschaften verändert wird, die durch andere Einflüsse als die, die in dieser Unterrichtsstichprobe wirksam sind, beeinflusst werden.« Professor Gage hat hier eine wichtige Frage angeschnitten, indem er darauf hinweist, daß die Gültigkeit der Ergebnisse solcher Untersuchungen nicht in erster Linie von dem Umfang der Unterrichtsstichprobe, sondern von der Gültigkeit und *Sensitivität* der erhaltenen Kriteriumswerte abhängig ist.